Gaussian Beam Optics

Dathon Golish
OPTI 521
Tutorial

University of Arizona
520-626-9230
dgolish@u.arizona.edu
December 8th, 2006

I. Introduction

In your typical optics lab, most of us, as scientists or engineers, are familiar with geometrical optics, which assumes that the wavelength of the light approaches zero. When constructing a formalism for light propagation and optical analysis, this assumption allows us to ignore any effects resulting from diffraction and greatly simplifies most calculations. This approximation – that the wavelength is very small compared to the system dimensions – is a very good one for optical wavelengths. In general, though it is dependent on the system dimensions, this assumption holds well for all wavelengths shorter than the mid-IR (~25 microns). However, when one goes past this point, the assumption begins to break down. Optical design in submillimeter astronomy has as much to do with radio optics as it does with traditional optics. At these wavelengths (~1 mm), the system dimensions no longer dominate and diffraction begins to play a large role. In this paper I will discuss and illustrate through real examples, how this change of regime affects optical designs and what considerations that can bring to our optomechanical designs.

II. Theory

The primary difference in optical design, when transitioning between geometrical and Gaussian beam optics, is the impact of wavelength on diffraction effects. At optical wavelengths, the entrance pupil for standard telescopes (ranging from a few inches to several meters) can be anywhere from one hundred thousand to tens of millions of wavelengths across. In this regime, diffraction is clearly negligible and geometric optics work extremely well. However, if you look at this same range of telescopes at a wavelength of one millimeter, you find that the entrance pupil can be anywhere from tens of thousands to as few as one hundred wavelengths across. Diffraction can no longer be ignored for these frequencies and system scales. At each point throughout the system, the beam bundle can and must be treated as a Gaussian beam, down to and including the focus. This concept is not truly so foreign to an engineer in the optical regime, as the

Gaussian focus of a radio telescope has the very same shape as an Airy disk, including a first minimum at ~1.22 λf/D. The only difference is that for optical wavelengths, the central spot is perhaps 50 nm across. At submillimeter wavelengths, this central spot grows to 0.1 mm and will clearly have diffraction effects when focused on a detector that is only 0.5 mm across. Only very well designed optical wavelength systems can be treated as diffraction limited – and then only in the sense that the optical surfaces are designed and manufactured to this degree of accuracy; but in radio and submillimeter astronomy, the system is *always* diffraction limited.

Detector technology can also be very different in the submillimeter regime. In particular, the research I do is in the coherent branch of submillimeter astronomy. Incoherent detectors, such as bolometers, are used as simple temperature detectors, much like your standard CCD optical system. Coherent detectors measure both the amplitude and the phase of the incoming wave. This allows spectroscopy to be done on the astronomical signals, something one cannot do in a straightforward manner with incoherent detectors. Coherent detection, however, is done in waveguide. This requires coupling of free space light waves to waveguide. The standard method of doing this uses feedhorns which capture the light as a single-moded Gaussian and send it down the waveguide to the detector. The apertures for these feedhorns are typically on the order of a few wavelengths across and are designed to transmit (and therefore receive) Gaussian beams. Therefore, any optical analysis that will be performed on them must fundamentally be done in the Gaussian regime. In order to create a complete optical system, one must couple the incoming light which is focused by the telescope to the power pattern of the feedhorn.

Gaussian beam formalism is a direct solution to the Helmholtz wave equation (shown in this case for the electric field, but equally applicable to the magnetic field).

$$\left(\nabla^2 + k^2\right)\Psi = \frac{\delta^2 E}{\delta x^2} + \frac{\delta^2 E}{\delta y^2} + \frac{\delta^2 E}{\delta z^2} + k^2 E = 0 \qquad (1)$$

A Gaussian field does not have infinite extent like a plane wave in geometrical optics, it varies as a Gaussian in the directions perpendicular to propagation. This formalism assumes that the electric and magnetic fields are perpendicular to each other and to the direction of propagation. If we let the direction of propagation be along the positive *z* direction, we assume the following form for the electric field, ignoring time dependence.

$$E(x,y,z) = u(x,y,z)e^{-jkz} \qquad (2)$$

Plugging equation (2) into equation (1), we obtain the **reduced wave equation**.

$$\frac{\delta^2 u}{\delta x^2} + \frac{\delta^2 u}{\delta y^2} + \frac{\delta^2 u}{\delta z^2} - 2jk\frac{\delta u}{dz} = 0 \tag{3}$$

At this point we must make an approximation to simplify the solutions to the wave equation. If we assume that the beams are basically paraxial (which, in practice, includes beams that are mostly confined to within 30 degrees of the *z* axis), we can make two approximations. First, that the variation of the field in the propagation direction is small over distances on the order of a wavelength. Second, that the variation of the field in the direction of propagation will also be small compared to the variation in the directions perpendicular to it. These approximations allow us to disregard the third term in equation (3) as it is negligible compared to both the first two terms and the last term. Adopting this, we obtain the **paraxial wave equation**.

$$\frac{\delta^2 u}{\delta x^2} + \frac{\delta^2 u}{\delta y^2} - 2jk\frac{\delta u}{dz} = 0 \tag{4}$$

The solutions to the paraxial wave equation are the Gaussian beam modes. These solutions, as well as much of Gaussian beam formalism, are well-known to those familiar with laser design and engineering. Considered in cylindrical coordinates, and assuming axial symmetry, the paraxial equation becomes

$$\frac{\delta^2 u}{\delta r^2} + \frac{1}{r}\frac{\delta u}{\delta r} - 2jk\frac{\delta u}{dz} = 0 \tag{5}$$

We assume a basic form for the wave equation of

$$u(x,z) = A(z) \cdot e^{-jkr^2/2q(z)} \tag{5}$$

where $q(z)$ is the complex Gaussian beam parameter. This solution can be more fully developed and $q(z)$ can be solved in terms of R, the radius of curvature of the wavefront, and w, the beam radius, which is the radius at which the field is equation to $1/e$ of its on-axis value. The beam radius is at its smallest (i.e. the focus) when $z=0$ and is called the **beam waist radius**, denoted w_0. A detailed examination of this derivation can be found in Goldsmith[2]. For the purposes of this paper, however, I will skip it and reveal the important elements, which are the solutions for R and w as a function of distance along the axis of propagation. Schematic diagrams of Gaussian beam propagation can be seen in Figure 1, taken from Goldsmith[2].

$$R = z + \frac{1}{z}\left(\frac{\pi {w_0}^2}{\lambda}\right)^2 \tag{5}$$

$$w = w_0 \sqrt{1 + \left(\frac{\lambda z}{\pi w_0^{\ 2}} \right)^2} \tag{6}$$

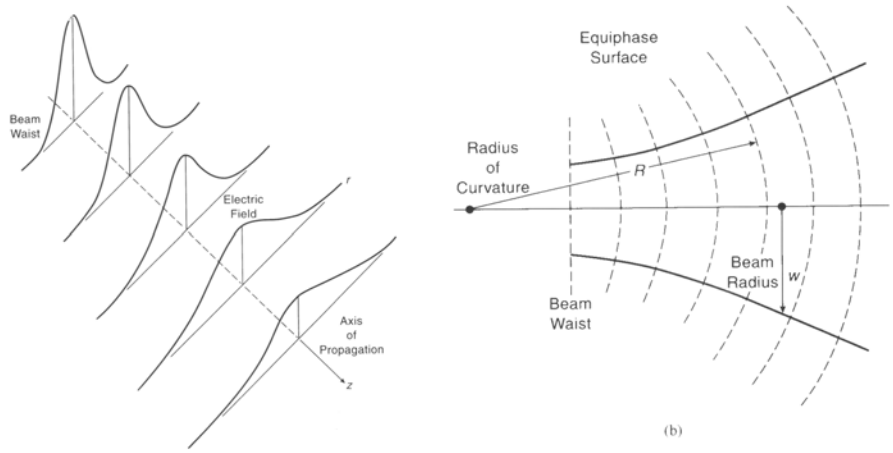

Figure 1: Gaussian Beam Propagation, taken from Goldsmith[2]

One can see that if we return to the geometric limit, as λ goes to 0, the radius of curvature of the wavefront goes to infinity, i.e. a plane wave. However, these equations must be used when designing any diffraction-limited optical system. This becomes particularly important when considering beam sizes. As a general rule, programs such as Zemax or Code V still perform fairly well in the submillimeter regime for basic optical design – e.g. optimizing lens and mirror surfaces. Though I will not discuss it here, it is most often the case that ray tracing programs can be used for the optical design of a submillimeter system. But, more rigorous programs such as Breault Research Organization's ASAP must be used to analyze the performance of these systems by taking into account diffraction effects.

However, the standard ray tracing programs do not manage Gaussian beam sizes well. If one designs an optical system entirely in Zemax and does not consider Gaussian beam sizes, vignetting will become a significant problem because the energy distribution spreads out much further than geometrical optics would predict. This matters not only because energy is lost around the edges of the optics, but also because, since we are in the diffraction regime, if we have a beam that is larger than a given aperture, we will introduce even more diffraction effects. These effects can lead to ripples and distortions in the Gaussian beam profile. These distortions reduce the match to the feedhorns, which couple most efficiently to a pure Gaussian.

Case Study #1 – PoleSTAR LO Beam Splitter

A prime example of this effect is illustrated by the Local Oscillator (LO) beam splitter designed for PoleSTAR. PoleSTAR is a 4-pixel 810 GHz array receiver designed by the Steward Observatory Radio Astronomy Lab (SORAL), for the Antarctic Submillimeter Telescope and Remote Observatory (AST/RO) formerly in operation at the South Pole. An AutoCAD image of PoleSTAR can be seen in Figure 2. The four PoleSTAR superheterodyne (coherent) pixels are arranged in a 2x2 grid (shown in Figure 2 as a single beam bundle), with 30 mm between adjacent pixels. Another aspect of coherent submillimeter detection is the need for a local oscillator source. Current computer technology can not process an 810 GHz sky signal in a spectrometer. Instead, the sky signal is beat against another reference signal (the LO) in the detector (or mixer). Basic interference theory tells us that two frequencies will be produced by this combination - a sum and a difference of the two initial waves. The LO is designed to have a frequency very close to the sky frequency, such that the difference frequency, or Intermediate Frequency (IF), will be around 5 GHz, which can be processed by a spectrometer. Though these signals can sometimes be injected in waveguide within the mixer, they are most often injected quasi-optically, usually with a beam splitter, before the sky signal reaches the detector. This was the case in PoleSTAR; but because it was an array receiver, the LO signal had to be distributed to the four pixels. This seems like a simple problem, as the signal can just be divided up by 2 mirrors and 2 beam splitters, as shown in Figure 3.

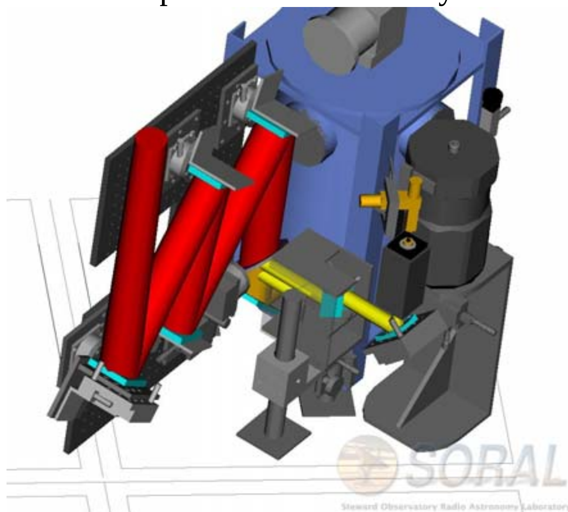

Figure 2: PoleSTAR on AST/RO

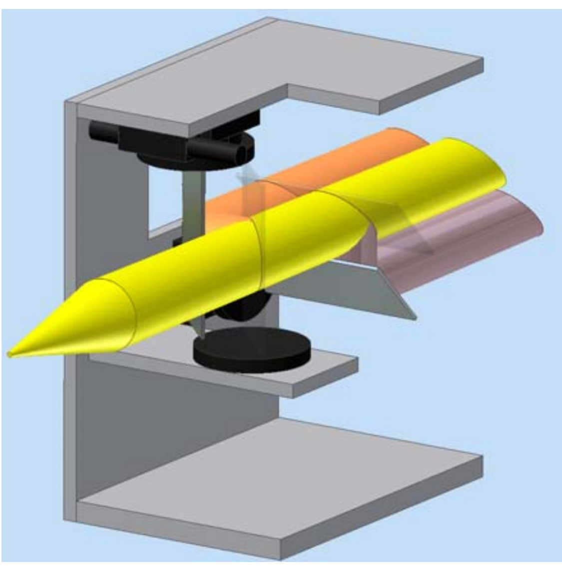

Figure 3: PoleSTAR LO Splitter

However, in practice, because of Gaussian beam effects, this problem becomes much more complicated. In radio optics, we often think of receivers as transmitters in order to couple incoming light to the detectors. As standard practice, we place plano-convex lenses one focal length in front of feedhorns, which emit very broad beams, to "collimate" the beam. However, because of diffraction effects, these beams are not truly collimated and spread out like Gaussian beams. If one ignored this fact, the beams, which would be the size of the lenses (27 mm in this case) would fit well through the LO splitter. However, because of the Gaussian expansion, these beams only fit to about the 95% level. Though this may seem insignificant, this can cause significant distortion in the Gaussian beam profile. Furthermore, there is now 5% of the LO power reflecting off errant surfaces which can cause interference or standing waves in the LO path. LO coupling is known to be very sensitive and small variations in the power that couples with the sky signal can complete destabilize mixing. Clearly in this case diffraction effects must be considered and accounted for. Unfortunately, due to the fixed spacing of the pixels, these effects could only be managed rather than eliminated. Nonetheless, an awareness of Gaussian beam effects is crucial.

III. Optomechanical Considerations

A. Tolerances

In general, engineers in the submillimeter regime have a much easier task when it comes to tolerancing their optical systems. Because the wavelengths are so much larger, wavelength-dependant factors such as surface alignment are usually much more lenient. A typical optical system might have linear and angular positioning tolerances that are measured in microns or milliradians. On the other hand, a system with a wavelength 1000 times greater will be much more forgiving.

Case Study #2 – SuperCam Tolerancing

SuperCam is another array receiver currently being developed in SORAL. SuperCam is a 64-pixel 345 GHz heterodyne array receiver designed for the Heinrich Hertz Telescope (HHT) on Mt. Graham in southern Arizona. An AutoCAD image of SuperCam in the apex room behind the 10 m HHT dish can be seen in Figure 4.

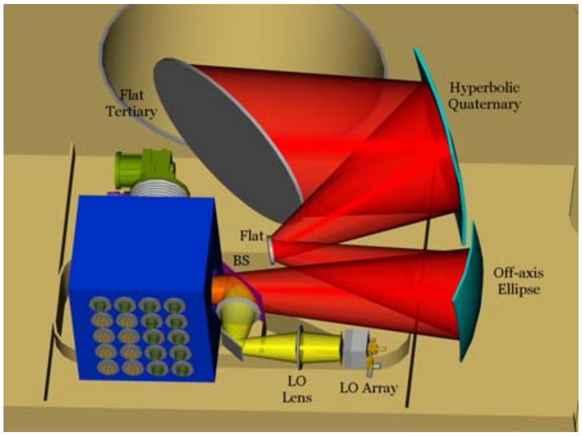

Figure 4: SuperCam in the Apex Room of the HHT

SuperCam's optical system basically consists off an off-axis hyperbola and an off-axis ellipse which change the f/13.8 telescope beam to an f/5 to match the feedhorns of the detector, which are inside the large blue cryostat. Nominally, two off-axis mirrors would produce a considerable challenge in alignment and optomechanical stability. However, when the tolerances are calculated in Zemax, which is still well suited for this task, notable differences are revealed. A collection of Zemax's results can be seen in Tables 1 and 2.

Table 1: SuperCam Linear Shift Mirror Tolerances

Optic	Shift (along axis) (mm)			
	-X	+X	-Y	+Y
Flat Tertiary	†	†	†	†
Hyperbola	-15.8	15.8	-21.5	17.4
Flat Fold	†	†	†	†
Ellipse	-16	16	-18.3	14.4

†Shift has no effect.

Table 2: SuperCam Angular Shift Mirror Tolerances

Optic	Rotation (around axis) (deg)					
	-X	+X	-Y	+Y	-Z	+Z
Flat Tertiary	-.72	3.2	-1.5	1.5	*	*
Hyperbola	-1.7	1.3	-1.7	1.7	-21.5	21.5
Flat Fold	-2.6	3.0	-3.1	3.1	*	*
Ellipse	-1.4	1.3	-1.3	1.3	-21.5	21.5

*Shift has no effect.

Clearly, SuperCam's sensitivity to positional errors is much more lenient than a typical optical system. With positional errors over half an inch and angular errors between one and three degrees, the system is much easier to align and keep stable during observations. In this case, operating at such large wavelengths is actually a benefit, as many of the normal optomechanical issues become irrelevant.

Lenient tolerances in the submillimeter regime become very convenient when considering the fact that one cannot align the optics visually. Even in the infrared there are IR-sensitive cards you can use to track a beam path, but in the submillimeter there is no way to visually track the beam path. Moreover, the detectors only work at cryogenic temperatures (as discussed in section D) so a quick measurement cannot be easily taken to determine alignment either. Typically, one has to design a laser alignment module that is mounted to simulate the telescope beam and align the optics. However, submillimeter dielectrics such as high density polyethylene do not transmit at optical wavelengths, so this method still only works for reflective systems. Therefore, submillimeter optical systems have to start very well aligned without user intervention or it will be too difficult to find the proper alignment.

B. Optical Manufacturing

At long wavelengths, the manufacturing of optical elements also becomes much easier. As a general rule, we prefer to make optics with surface roughnesses around $\lambda/10$ rms, sometimes lower depending on the level of accuracy required. However, because our wavelength is 1000 times larger, our rms surface roughness is correspondingly larger. For example, at 300 GHz (1 mm), this implies only a 0.1 mm rms surface roughness. This makes the creation of flat mirrors much simpler for the manufacturer. An average precision lapping machine found in your typical optics shop will often be sufficient. Additionally, the tolerances on radii of curvature can be relatively lenient for the same reasons. We can normally make our optical elements on standard CNC machining equipment.

Case Study #3 – SuperCam Optical Manufacturing

Returning again to the SuperCam array receiver, a tolerancing was also done on the optical components to determine their manufacturing requirements, shown in Table 3. Recalling Figure 4, the system consists of two flats and two curved mirrors.

Table 3: Surface Quality Tolerances

Optic	Surface Quality Quantity	Value	
Flat Tertiary	Curvature Error in Fringes	-.43	.57
Flat Tertiary	Irregularity in Fringes	-.28	.31
Hyperbola	Radius of Curvature in mm (1692 mm Nominal)	-54	84.5
Hyperbola	Irregularity in Fringes	-1.67	1.4
Flat Fold	Curvature Error in Fringes	-3.2	2.9
Flat Fold	Irregularity in Fringes	-2.1	3
Ellipse	Radius of Curvature in mm (772 mm Nominal)	-16.7	25.8
Ellipse	Irregularity in Fringes	-4.6	5.6

These surface quality quantities, except radius of curvature, are all measured in fringes, which at this wavelength (870 microns) correspond to only 0.435 mm. Producing a maximum curvature error of 0.2 mm for a flat mirror is not a difficult task. Though its large size – around 0.7 m – will present a problem, this is not really a function of the frequency regime. Furthermore, the radii of curvature must be held to somewhere around 3% of their radius, which is considered base precision for typical machining. These mirrors will be manufactured on a computer-controlled milling machine in the Steward Observatory shop and do not require a specialty machine shop.

C. Vibration & Stability

One area of optomechanics where submillimeter astronomy is not as lenient is in system stability. As described in Case Study #1, the mixing of sky signals and local oscillator signals can be very sensitive to variation. If an LO signal is unstable, whether because of the oscillator itself or vibrations along the signal path, mixing can easily break down. This problem becomes especially significant as we approach the terahertz end of our spectrum, as signal generation technology becomes less reliable.

Case Study #4 – TREND

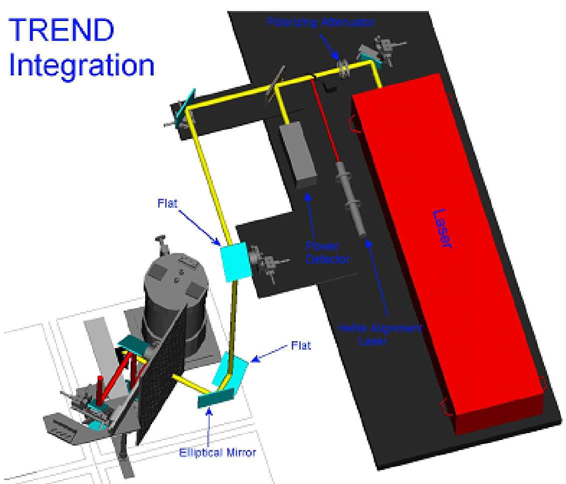

Figure 5: TREND FIR Laser LO

TeraHertz REceiver with an NbN Device (TREND) was a 1.5 THz single-pixel receiver designed by the University of Massachusetts for operation on AST/RO at the South Pole. The optics, however, were done in SORAL. Because at the time no other LO devices were cheaply available at 1.5 THz, TREND used a far infrared laser to create its signal. This laser was somewhat unstable and made observations difficult. In addition, due to its size, the laser, along with most of its optics, was mounted in the ceiling above the receiver pallets. This caused significant vibration problems in the optical path and often completely destabilized mixing in the receiver. Had more care been paid to the vibrational characteristics of the LO optical system, these problems might have been somewhat mitigated. This problem occurs repeatedly in submillimeter receivers, often because the LO must be injected from a direction perpendicular to the optical table and is therefore often on a long lever arm subject to significant vibrations.

D. Cryogenic Temperatures

Another aspect of submillimeter detector technology is the necessity of cryogenic temperatures. It is often the case that detector technology in the optical regime works better at colder temperatures, in part because of the reduction in dark current, but it is not a necessity for most CCDs to function. However, the most common types of heterodyne detectors are SIS devices, or Superconductor-Insulator-Superconductor devices. In the SIS device, Cooper pairs (pairs of bound electrons) are allowed to travel between superconductors because they have the same energy levels. However, if one biases this gap with an applied voltage, the energy levels separate and the Cooper pairs are blocked from crossing the insulator. Thus, only when an incoming photon supplies the appropriate amount of energy are Cooper pairs permitted to travel across the junction to the empty energy levels of the other superconductor. Hence, we can convert an electrical current to a photon count. However, this device obviously will only work when the superconductors are kept superconducting, which for most materials implies at least liquid helium (LHe) temperatures (4K); if not even lower (as low as 0.3 K using liquid Helium-3). A schematic of the device can be seen in Figure 5, supplied by Craig Kulesa[3] from the SORAL website.

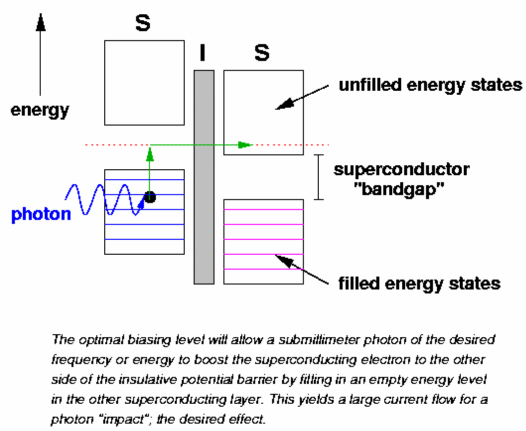

Figure 6: SIS Device Schematic

As a result, all of our detectors are designed in cryostats keeping them at LHe temperatures. This naturally means that all mounting elements for the detectors, as well as the corresponding electronics (which also appreciate the lower temperatures as they are less noisy) are kept at 4K or lower. We most often use oxygen-free copper at the 4K stage because its excellent thermal conductivity transmits the heat away from the mixers to the LHe pot efficiently. But designing mixer mounts that will experience almost a 300 Kelvin shift in temperature requires special attention, especially because of copper's moderately high coefficient of thermal expansion (~16.9 x 10^{-6}/K)[5].

Case Study #5 – SuperCam Mixer Block

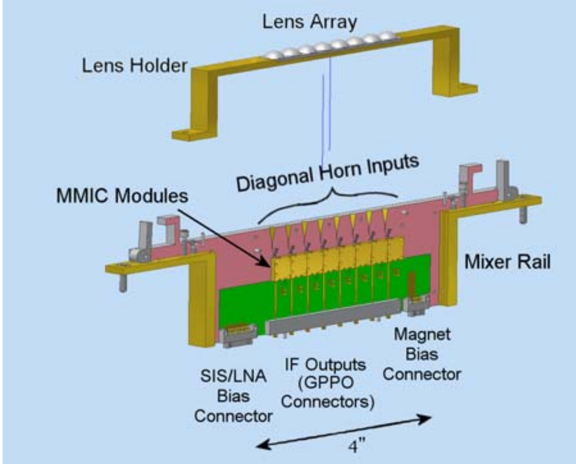

Figure 7: SuperCam 1x8 Mixer Block

Again examining the SuperCam receiver, Figure 6 shows an Inventor drawing of a 1x8 SuperCam mixer block. SuperCam's 8x8 array will be composed of eight 1x8 "cards." This was done to simplify assembly, as the mixers require around 10 wires each. The mixer design includes an IF board (shown as green) which transmits the downconverted sky signals along a microstrip line to SMA-type coaxial connectors, as well as the many bias signals required to control each mixer. This board has a Rogers Corporation TMM 3 substrate, which is quoted[4] as having a CTE of 37 x 10^{-6}/K. Because of the length of the board (to accommodate 8 pixels, each 11 mm apart from its neighbor), the thermal cycle causes a differential contraction of around 0.6 mm. Where as we might normally glue a board like this down, a differential thermal contraction of over half a millimeter might place too much stress on the board and shatter it. Therefore we had to design the board to have oversized mounting holes (and an oversized pocket) capable of accommodating this shift. This is just one example of how working at cryogenic temperatures can cause thermal contractions that pose a significant optomechanical problem. Clearly, thermal issues must be carefully managed in heterodyne submillimeter receiver systems.

V. Conclusion

Though this is of course just a sampling, this report is meant to introduce a number of optomechanical considerations that are introduced when working in the submillimeter regime. Working at these wavelengths can be very different than what most optical engineers are used to. But as FIR technology pushes to longer wavelengths and more applications are found for terahertz technology, these issues will become more commonplace and more important to understand.

References:

1. Burge, Jim. "Introductory Opto-Mechanical Engineering." 2006. Optical Sciences Center, University of Arizona. http://www.optics.arizona.edu/optomech/Fall06/

2. Goldsmith, Paul F. Quasioptical Systems. New York: IEEE Press. 1998.

3. Kulesa, Craig. "SORAL Website." 1998. Steward Observatory, University of Arizona. http://soral.as.arizona.edu/overview.html

4. Rogers Corporation. "TMM Laminates Data Sheet." http://www.rogerscorporation.com/mwu/pdf/tmmdata.pdf

5. Yoder Jr., Paul R. Mounting Optics in Optical Instruments. Bellingham: SPIE. 2002.

6. Walker, Chris. *Class Notes from ASTR 585*. 2005. Steward Observatory, University of Arizona.